

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/575, 374  
Source: IFW0  
Date Processed by STIC: 02/26/2007

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 02/26/2007

PATENT APPLICATION: US/10/575,374

TIME: 14:31:48

Input Set : N:\efs\02\_26\_07\10575374\_efs\sequence.txt

Output Set: N:\CRF4\02262007\J575374.raw

```

3 <110> APPLICANT: Ecole Polytechnique Federale de Lausanne (EPFL)
5 <120> TITLE OF INVENTION: Method for identification of suitable fragmentation sites in
a re-
6     porter protein
8 <130> FILE REFERENCE: PEPF001WO
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/575,374
C--> 10 <141> CURRENT FILING DATE: 2006-04-10
10 <150> PRIOR APPLICATION NUMBER: US 34,404 JM-213
11 <151> PRIOR FILING DATE: 2003-10-09
13 <160> NUMBER OF SEQ ID NOS: 66
15 <170> SOFTWARE: PatentIn version 3.1
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 672
19 <212> TYPE: DNA
20 <213> ORGANISM: Saccharomyces cerevisiae
22 <220> FEATURE:
23 <221> NAME/KEY: modified_base
24 <222> LOCATION: (186)..(186)
25 <223> OTHER INFORMATION: silent point mutation introduced to generate HindIII
restriction
26     site
29 <300> PUBLICATION INFORMATION:
30 <308> DATABASE ACCESSION NO: NCBI / NC_001136
31 <309> DATABASE ENTRY DATE: 2004-08-30
33 <400> SEQUENCE: 1
34 atgtctgtta ttaatttcac aggtagttct ggtccattgg tgaaagtttg cggcttgcag      60
36 agcacagagg ccgcagaatg tgctctagat tccgatgctg acttgctggg tattatatgt      120
38 gtgccaata gaaagagaac aattgaccgc gttattgcaa ggaaaatttc aagtcttgta      180
40 aaagcatata aaaatagttc aggcactccg aaataacttg ttggcgtggt tcgtaatcaa      240
42 cctaaggagg atgttttggc tctggtcaat gattacggca ttgatatcgt ccaactgcac      300
44 ggagatgagt cgtggcaaga ataccaagag ttcctcggtt tgccagttat taaaagactc      360
46 gtatttccaa aagactgcaa catactactc agtgcagctt cacagaaacc tcattcgttt      420
48 attcccttgt ttgattcaga agcaggtggg acaggtgaac ttttggattg gaactcgatt      480
50 tctgactggg ttggaaggca agagagcccc gagagcttac attttatgtt agctggtgga      540
52 ctgacgccag aaaatggttg tgatgcgctt agattaaatg gcgttatttg tggtgatgta      600
54 agcggaggtg tggagacaaa tgggtgtaaaa gactctaaca aaatagcaaa ttctgtaaaa      660
56 aatgctaaga aa
59 <210> SEQ ID NO: 2
60 <211> LENGTH: 224
61 <212> TYPE: PRT
62 <213> ORGANISM: Saccharomyces cerevisiae
64 <300> PUBLICATION INFORMATION:
65 <308> DATABASE ACCESSION NO: NCBI / NC_001136
66 <309> DATABASE ENTRY DATE: 2004-08-30

```

68 <400> SEQUENCE: 2

## RAW SEQUENCE LISTING

DATE: 02/26/2007

PATENT APPLICATION: US/10/575,374

TIME: 14:31:48

Input Set : N:\efs\02\_26\_07\10575374\_efs\sequence.txt

Output Set: N:\CRF4\02262007\J575374.raw

```

70 Met Ser Val Ile Asn Phe Thr Gly Ser Ser Gly Pro Leu Val Lys Val
71 1 5 10 15
74 Cys Gly Leu Gln Ser Thr Glu Ala Ala Glu Cys Ala Leu Asp Ser Asp
75 20 25 30
78 Ala Asp Leu Leu Gly Ile Ile Cys Val Pro Asn Arg Lys Arg Thr Ile
79 35 40 45
82 Asp Pro Val Ile Ala Arg Lys Ile Ser Ser Leu Val Lys Ala Tyr Lys
83 50 55 60
86 Asn Ser Ser Gly Thr Pro Lys Tyr Leu Val Gly Val Phe Arg Asn Gln
87 65 70 75 80
90 Pro Lys Glu Asp Val Leu Ala Leu Val Asn Asp Tyr Gly Ile Asp Ile
91 85 90 95
94 Val Gln Leu His Gly Asp Glu Ser Trp Gln Glu Tyr Gln Glu Phe Leu
95 100 105 110
98 Gly Leu Pro Val Ile Lys Arg Leu Val Phe Pro Lys Asp Cys Asn Ile
99 115 120 125
102 Leu Leu Ser Ala Ala Ser Gln Lys Pro His Ser Phe Ile Pro Leu Phe
103 130 135 140
106 Asp Ser Glu Ala Gly Gly Thr Gly Glu Leu Leu Asp Trp Asn Ser Ile
107 145 150 155 160
110 Ser Asp Trp Val Gly Arg Gln Glu Ser Pro Glu Ser Leu His Phe Met
111 165 170 175
114 Leu Ala Gly Gly Leu Thr Pro Glu Asn Val Gly Asp Ala Leu Arg Leu
115 180 185 190
118 Asn Gly Val Ile Gly Val Asp Val Ser Gly Gly Val Glu Thr Asn Gly
119 195 200 205
122 Val Lys Asp Ser Asn Lys Ile Ala Asn Phe Val Lys Asn Ala Lys Lys
123 210 215 220
126 <210> SEQ ID NO: 3
127 <211> LENGTH: 132
128 <212> TYPE: DNA
129 <213> ORGANISM: Saccharomyces cerevisiae
131 <220> FEATURE:
132 <221> NAME/KEY: modified_base
133 <222> LOCATION: (22)..(22)
134 <223> OTHER INFORMATION: point mutation
137 <400> SEQUENCE: 3
138 atgtctgtta ttaatttcac atgtagttct ggtccattgg tgaaagtttg cggcttgcag 60
140 agcacagagg ccgcagaatg tgctctagat tccgatgctg acttgctggg tattatatgt 120
142 gtgcccaata ga 132
145 <210> SEQ ID NO: 4
146 <211> LENGTH: 44
147 <212> TYPE: PRT
148 <213> ORGANISM: Saccharomyces cerevisiae
150 <400> SEQUENCE: 4
152 Met Ser Val Ile Asn Phe Thr Cys Ser Ser Gly Pro Leu Val Lys Val
153 1 5 10 15
156 Cys Gly Leu Gln Ser Thr Glu Ala Ala Glu Cys Ala Leu Asp Ser Asp
157 20 25 30

```

## RAW SEQUENCE LISTING

DATE: 02/26/2007

PATENT APPLICATION: US/10/575,374

TIME: 14:31:48

Input Set : N:\efs\02\_26\_07\10575374\_efs\sequence.txt

Output Set: N:\CRF4\02262007\J575374.raw

```

160 Ala Asp Leu Leu Gly Ile Ile Cys Val Pro Asn Arg
161          35          40
164 <210> SEQ ID NO: 5
165 <211> LENGTH: 540
166 <212> TYPE: DNA
167 <213> ORGANISM: Saccharomyces cerevisiae
169 <400> SEQUENCE: 5
170 aagagaacaa ttgacccggt tattgcaagg aaaatttcaa gtcttgtaaa agcatataaa      60
172 aatagttcag gcactccgaa atacttggtt ggcgtgtttc gtaatcaacc taaggaggat      120
174 gttttggctc tggtaaatga ttacggcatt gatatcgccc aactgcacgg agatgagtcg      180
176 tggcaagaat accaagagtt cctcggtttg ccagttatta aaagactcgt atttccaaaa      240
178 gactgcaaca tactactcag tgcagcttca cagaaacctc attcgtttat tcccttgttt      300
180 gattcagaag caggtgggac aggtgaactt ttggattgga actcgatttc tgactggggt      360
182 ggaaggcaag agagccccga gagcttacat tttatgtag ctggtggact gacgccagaa      420
184 aatgttggtg atgcgcttag attaaatggc gttattggtg ttgatgtaag cggaggtgtg      480
186 gagacaaatg gtgtaaaaga ctctaacaaa atagcaaatt tcgtcaaaaa tgctaagaaa      540
189 <210> SEQ ID NO: 6
190 <211> LENGTH: 180
191 <212> TYPE: PRT
192 <213> ORGANISM: Saccharomyces cerevisiae
194 <400> SEQUENCE: 6
196 Lys Arg Thr Ile Asp Pro Val Ile Ala Arg Lys Ile Ser Ser Leu Val
197 1          5          10          15
200 Lys Ala Tyr Lys Asn Ser Ser Gly Thr Pro Lys Tyr Leu Val Gly Val
201          20          25          30
204 Phe Arg Asn Gln Pro Lys Glu Asp Val Leu Ala Leu Val Asn Asp Tyr
205          35          40          45
208 Gly Ile Asp Ile Val Gln Leu His Gly Asp Glu Ser Trp Gln Glu Tyr
209          50          55          60
212 Gln Glu Phe Leu Gly Leu Pro Val Ile Lys Arg Leu Val Phe Pro Lys
213 65          70          75          80
216 Asp Cys Asn Ile Leu Leu Ser Ala Ala Ser Gln Lys Pro His Ser Phe
217          85          90          95
220 Ile Pro Leu Phe Asp Ser Glu Ala Gly Gly Thr Gly Glu Leu Leu Asp
221          100         105         110
224 Trp Asn Ser Ile Ser Asp Trp Val Gly Arg Gln Glu Ser Pro Glu Ser
225          115         120         125
228 Leu His Phe Met Leu Ala Gly Gly Leu Thr Pro Glu Asn Val Gly Asp
229          130         135         140
232 Ala Leu Arg Leu Asn Gly Val Ile Gly Val Asp Val Ser Gly Gly Val
233 145         150         155         160
236 Glu Thr Asn Gly Val Lys Asp Ser Asn Lys Ile Ala Asn Phe Val Lys
237          165         170         175
240 Asn Ala Lys Lys
241          180
244 <210> SEQ ID NO: 7
245 <211> LENGTH: 159
246 <212> TYPE: DNA
247 <213> ORGANISM: Saccharomyces cerevisiae

```

## RAW SEQUENCE LISTING

DATE: 02/26/2007

PATENT APPLICATION: US/10/575,374

TIME: 14:31:48

Input Set : N:\efs\02\_26\_07\10575374\_efs\sequence.txt

Output Set: N:\CRF4\02262007\J575374.raw

```

249 <400> SEQUENCE: 7
250 atgtctgtta ttaatttcac aggtagttct ggtccattgg tgaaagtttg cggcttgca 60
252 agcacagagg ccgcagaatg tgctctagat tccgatgctg acttgctggg tattatatgt 120
254 gtgcccaata gaaagagaac aattgacccg gttattgca 159
257 <210> SEQ ID NO: 8
258 <211> LENGTH: 53
259 <212> TYPE: PRT
260 <213> ORGANISM: Saccharomyces cerevisiae
262 <400> SEQUENCE: 8
264 Met Ser Val Ile Asn Phe Thr Gly Ser Ser Gly Pro Leu Val Lys Val
265 1 5 10 15
268 Cys Gly Leu Gln Ser Thr Glu Ala Ala Glu Cys Ala Leu Asp Ser Asp
269 20 25 30
272 Ala Asp Leu Leu Gly Ile Ile Cys Val Pro Asn Arg Lys Arg Thr Ile
273 35 40 45
276 Asp Pro Val Ile Ala
277 50
280 <210> SEQ ID NO: 9
281 <211> LENGTH: 516
282 <212> TYPE: DNA
283 <213> ORGANISM: Saccharomyces cerevisiae
285 <400> SEQUENCE: 9
286 gcaaggaaaa tttcaagtct tgtaaaagct tataaaaata gttcaggcac tccgaaatac 60
288 ttggttggcg tgtttcgtaa tcaacctaa gaggatgttt tggctctggg caatgattac 120
290 ggcattgata tcgtccaact gcacggagat ggtcgtggc aagaatacca agagttcctc 180
292 ggtttgccag ttattaaaag actcgtattt ccaaaagact gcaacatact actcagtgca 240
294 gcttcacaga aacctcattc gtttattccc ttgtttgatt cagaagcagg tgggacagg 300
296 gaacttttgg attggaactc gatttctgac tgggttgaa ggcaagagag ccccgagagc 360
298 ttacatttta tgtagctgg tggactgacg ccagaaaatg ttggtgatgc gcttagatta 420
300 aatggcgtaa ttggtgttga tgtaagcggg ggtgtggaga caaatggtgt aaaagactct 480
302 aacaaaatag caaatctcgt caaaaatgct aagaaa 516
305 <210> SEQ ID NO: 10
306 <211> LENGTH: 172
307 <212> TYPE: PRT
308 <213> ORGANISM: Saccharomyces cerevisiae
310 <400> SEQUENCE: 10
312 Ala Arg Lys Ile Ser Ser Leu Val Lys Ala Tyr Lys Asn Ser Ser Gly
313 1 5 10 15
316 Thr Pro Lys Tyr Leu Val Gly Val Phe Arg Asn Gln Pro Lys Glu Asp
317 20 25 30
320 Val Leu Ala Leu Val Asn Asp Tyr Gly Ile Asp Ile Val Gln Leu His
321 35 40 45
324 Gly Asp Glu Ser Trp Gln Glu Tyr Gln Glu Phe Leu Gly Leu Pro Val
325 50 55 60
328 Ile Lys Arg Leu Val Phe Pro Lys Asp Cys Asn Ile Leu Leu Ser Ala
329 65 70 75 80
332 Ala Ser Gln Lys Pro His Ser Phe Ile Pro Leu Phe Asp Ser Glu Ala
333 85 90 95
336 Gly Gly Thr Gly Glu Leu Leu Asp Trp Asn Ser Ile Ser Asp Trp Val

```

## RAW SEQUENCE LISTING

DATE: 02/26/2007

PATENT APPLICATION: US/10/575,374

TIME: 14:31:48

Input Set : N:\efs\02\_26\_07\10575374\_efs\sequence.txt

Output Set: N:\CRF4\02262007\J575374.raw

```

337          100          105          110
340 Gly Arg Gln Glu Ser Pro Glu Ser Leu His Phe Met Leu Ala Gly Gly
341          115          120          125
344 Leu Thr Pro Glu Asn Val Gly Asp Ala Leu Arg Leu Asn Gly Val Ile
345          130          135          140
348 Gly Val Asp Val Ser Gly Gly Val Glu Thr Asn Gly Val Lys Asp Ser
349 145          150          155          160
352 Asn Lys Ile Ala Asn Phe Val Lys Asn Ala Lys Lys
353          165          170
356 <210> SEQ ID NO: 11
357 <211> LENGTH: 561
358 <212> TYPE: DNA
359 <213> ORGANISM: Saccharomyces cerevisiae
361 <400> SEQUENCE: 11
362 atgtctgtta ttaatttcac aggtagttct ggtccattgg tgaaagtttg cggcttgcag      60
364 agcacagagg ccgcagaatg tgctctagat tccgatgctg acttgctggg tattatatgt      120
366 gtgcccataa gaaagagaac aattgacccg gttattgcaa ggaaaatttc aagtcttgta      180
368 aaagcttata aaaatagttc aggcactccg aaatacttgg ttggcgtggt tcgtaatcaa      240
370 cctaaggagg atgttttggc tctggtcaat gattacggca ttgatatcgt ccaactgcac      300
372 ggagatgagt cgtggcaaga ataccaagag ttctctcggt tgccagttat taaaagactc      360
374 gtatttccaa aagactgcaa catactactc agtgcagctt cacagaaacc tcattcgttt      420
376 attcccttgt ttgattcaga agcaggtggg acaggtgaac ttttggattg gaactcgatt      480
378 tctgactggg ttggaaggca agagagcccc gagagcttac attttatgtt agctggtgga      540
380 ctgacgccag aaaatgttgg t                                     561
383 <210> SEQ ID NO: 12
384 <211> LENGTH: 187
385 <212> TYPE: PRT
386 <213> ORGANISM: Saccharomyces cerevisiae
388 <400> SEQUENCE: 12
390 Met Ser Val Ile Asn Phe Thr Gly Ser Ser Gly Pro Leu Val Lys Val
391 1          5          10          15
394 Cys Gly Leu Gln Ser Thr Glu Ala Ala Glu Cys Ala Leu Asp Ser Asp
395          20          25          30
398 Ala Asp Leu Leu Gly Ile Ile Cys Val Pro Asn Arg Lys Arg Thr Ile
399          35          40          45
402 Asp Pro Val Ile Ala Arg Lys Ile Ser Ser Leu Val Lys Ala Tyr Lys
403          50          55          60
406 Asn Ser Ser Gly Thr Pro Lys Tyr Leu Val Gly Val Phe Arg Asn Gln
407 65          70          75          80
410 Pro Lys Glu Asp Val Leu Ala Leu Val Asn Asp Tyr Gly Ile Asp Ile
411          85          90          95
414 Val Gln Leu His Gly Asp Glu Ser Trp Gln Glu Tyr Gln Glu Phe Leu
415          100          105          110
418 Gly Leu Pro Val Ile Lys Arg Leu Val Phe Pro Lys Asp Cys Asn Ile
419          115          120          125
422 Leu Leu Ser Ala Ala Ser Gln Lys Pro His Ser Phe Ile Pro Leu Phe
423          130          135          140
426 Asp Ser Glu Ala Gly Gly Thr Gly Glu Leu Leu Asp Trp Asn Ser Ile
427 145          150          155          160

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/575,374

DATE: 02/26/2007

TIME: 14:31:49

Input Set : N:\efs\02\_26\_07\10575374\_efs\sequence.txt

Output Set: N:\CRF4\02262007\J575374.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:480 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:15